UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/815,768	03/23/2001	John Kroeker	57622-036 (ELZ-1)	5839

7590 02/26/2004

Toby H. Kusmer
McDERMOTT, WILL & EMERY
28 State Street
Boston, MA 02109

EXAMINER	
LAO, TIM P	

ART UNIT	PAPER NUMBER
2655	8

DATE MAILED: 02/26/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO-90C (Rev. 10/03)

Office Action Summary	Application No.	Applicant(s)	
	09/815,768	KROEKER, JOHN	
	Examiner	**Art Unit**	
	Tim Lao	2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *23 March 2001*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-27* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-3,6-9,11-17,19,20,22,23, and 26* is/are rejected.

7)☒ Claim(s) *4,5,10,18,21,24,25, and 27* is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☒ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *2,3,4,5,6,7*.

4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

5) ☐ Notice of Informal Patent Application (PTO-152)

6) ☐ Other: _____.

DETAILED ACTION

Specification

1. The disclosure is objected to because of the following informalities: word such as "Novel" should be avoided in the Title. Appropriate correction is required.

2. The title of the invention is not descriptive. A new title is required that is clearly indicative of the invention to which the claims are directed. The following title is suggested: Speech recognition system and method for generating phonetic estimates.

Claim Rejections - 35 USC § 112

3. The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

4. Claims 6-9 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the enablement requirement. The claim(s) contains subject matter which was not described in the specification in such a way as to enable one skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention.

Regarding claim 6, it is not clear from the description in the specification how the first predetermined frequency range is substantially smaller than the second predetermined frequency range.

Regarding claim 7, it is not clear from the description in the specification how the first predetermined time span is substantially smaller than the second predetermined time span.

Regarding claim 8, it is not clear from the description in the specification how the second predetermined time span is large relative to the second predetermined frequency range.

Regarding claim 9, it is not clear from the description in the specification how the second predetermined frequency range is large relative to the second predetermined time span.

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

6. Claims 1-3, 11-17, 19, 20, 22, 23, and 26 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kroeker et al. (U.S. Patent 5,168,524) in view of Rabiner et al., (*Digital Processing of Speech Signals*, Prentice Hall, 1978).

Claim(s)	Kroeker et al. show:
1	A speech recognition system for transforming an acoustic signal into a stream of phonetic estimates, comprising: (see Abstract) a frequency analyzer (power spectrum analyzer, Fig.2: 18; Fig.3) for receiving the acoustic signal (speech signal s(t), Fig.3: 100) and producing as an output a short-time frequency representation (e_m, Fig.3: 114) of the acoustic signal; (col.7, ll.21-62) *{1. The discrete Fourier transform (DFT) of the finite length vector c_m **108**, i.e., the 128-point DFT vector d_m **110**, is a short-time Fourier transform of the acoustic signal. 2. The energy vector e_m **114** represents the energy spectrum in the frequency domain, i.e., the short-time frequency representation.}*

	a novelty processor (receptive field processor and adaptive field normalizer, Fig.2: 24, 26; Fig.6, 7) for receiving the short-time frequency representation (e.g., q_m, Fig.6: 130) of the acoustic signal, separating one or more background components (e.g., breathing noises, unvoiced phonemes, col.10, ll.23-27) of the representation from one or more region-of-interest components (e.g., the voice components, col.10, ll.15-16) of the representation, and producing a novelty output (X_n, Fig.7: 222) including the region of interest components (e.g., the voice components) of the representation according to one or more novelty parameters (parameters of the vector w_n, see Fig.7: 214 & col.10, ll.6-7; parameters p, t_n, and voice threshold, see Fig.7: 216 & col.10, ll.37-54); (see col.9, ll.59-68; col.10, ll.1-63) *{1. Fig.6 is a block diagram depicting the receptive field processor of Fig.2: 24; Fig.7 is a block diagram depicting the adaptive normalizer of Fig.2: 26.* *2. The vectors e_m **114**, f_m **116**, q_m **130**, and the matrix V_n **210**, all are short-time frequency representations of the acoustic signal through different stages of processing. For example, q_m, is the output of the inputs e_m and f_m through the intermediate steps of Fig.4 (see col.7, ll.63-68; col.8, ll.1-34). V_n is the result of processing q_m through the steps of Fig.6 and Fig.7: 206 and 208 (see col.9, ll.17-58).* *3. V_n **210** data correspond to a SPEECH signal segment with a significant presence of the voice components (col.9, ll.59-59-62). If the integrated energy, t_n, does not exceed the voice threshold value, e.g., 25, then the adaptive average vector x'_n **218**, which corresponds to the noise components in this case, is subtracted from V_n to produce the matrix X_n, i.e., the voice components (col.10, ll.17-29, 55-63; see Fig.7).}* an attention processor (energy detect processor, Fig.2: 22; Fig.5) for producing a gating signal (s_m, Fig.5: 134) according to one or more attention parameters (time parameter m and s_m values: 0, 1); a coincidence processor (receptive field nonlinear processor, Fig.2: 28; Fig.8, 9) for receiving the novelty output (X_n **222**) and producing a coincidence output (e.g., output of Fig.8: 228 & Fig.9: 234) that includes co-occurrences (e.g., correlations) between samples of the novelty output over time and frequency (Fig.8: 228 & Fig.9: 234; col.12, ll.1-13, 35-46), wherein the coincidence output is produced according to one or more coincidence parameters (e.g., delta time, j = 0…5; delta frequency, i = 1…20-Γ, Fig.8: 228); a vector pattern recognizer (Fig.2: 30, 32, 34; Fig.10, 11; col.13, ll.30-56) and a

	probability processor (Logarithm of the likelihood ratio processor, Fig.2: 42; Fig.14; col.16, ll.30-40) for receiving the coincidence output and producing a phonetic estimate stream (phoneme estimates, Fig.2: 46) representative of acoustic signal. *{The steps performed by the processor of Fig.2: 30, 32, and 34, i.e., the vector pattern recognizer, include concatenation of data into a vector and applying this vector to a speech element model so as to reduce the data of the vector to a set of speech element estimates. (col.13, ll.43-56)}* Kroeker et al. do not show: an attention processor for receiving the novelty output and producing a gating signal as a predetermined function of the novelty output. However, Rabiner et al. teach: receiving the novelty output (e.g., different speech samples (x(n+m), x(n+m+k), eq.4.33, p.146) and producing a gating signal (a windowed signal, $\hat{w}_1(m)$ and $\hat{w}_2(m)$, p.147, eq.4.35a, 4.35b) as a predetermined function of the novelty output (e.g., depending on the finite length N of the speech samples) according to one or more attention parameters (e.g., the time parameter m, the window values: 0, 1). (p.146-148) *{1. The gating signal is $\hat{w}_1(m)\hat{w}_2(m+k)$ when applying to the cross-correlation equation of 4.33, which can be written as eq.4.36. The gating signal is a function of time. Eq.4.36 is the cross-correlation function for two different finite length segments of speech (p.148, 2nd ¶).* *2. The finite length N, e.g., N = 6, would correspond to the time unit j = 0…5 of Fig.9: 234.}* It would have been obvious to a person of ordinary skill in the art at the time the invention was made to modify the speech recognition system of Kroeker et al. to include the windowing (i.e., gating signal) technique for calculating correlation functions (e.g, the modified short-time correlation function) as taught by Rabiner et al. in order to generate a selectively gated coincidence output. The benefit gain would be that more correlation peaks are displayed at the gated coincidence output (Rabiner et al., p.148, 2nd ¶, ll.6-7) which is useful for determining the periodicity of the speech signal.
Claim(s) 2	Kroeker et al. show:

	A speech recognition system according to claim 1, wherein the short-time frequency representation of the audio signal includes a series of consecutive time instances (e.g., 128-point, Fig.3: 110), each consecutive pair separated by a sampling interval (8KHz sampling interval, Fig.3: 110), and each of the time instances further includes a series of discrete Fourier transform (DFT) points ($c_{k,m}$, Fig.3: 108), such that the short-time frequency representation of the audio signal includes a series of DFT points ($d_{k,m}$, Fig.3: 108).
Claim(s) 3	Kroeker et al. show: A speech recognition system according to claim 2, wherein for each DFT point, the novelty processor (i) calculates a first average value (V_n Fig.7: 210; average by two in time; Fig.6: 204) across a first predetermined frequency range (0…20, Fig.6: 204) and a first predetermined time span (m…m-11, Fig.6: 204), (col.9, ll.17-25) *{The matrix U_n **206** is the result of averaging which becomes V_n **210** after the step of **208**.}* (ii) calculates a second average value (average over time, Fig.7: 212 and accumulative adaptive average, Fig.7: 216) across a second predetermined frequency range (0…20, Fig.7: 214) and a second predetermined time span (0…5, Fig.7: 212), (col.9, ll.59-68; col.10, ll.1-16) and *{The result of second average is the vector $x'_{k,n}$ **218**.}* (iii) subtracts (Fig.7: 220) the second average value ($x'_{k,n}$ **218**) from the first average value (V_n **210**) so as to produce the novelty output point (X_n **222**). (col.10, ll.55-63)
Claim(s) 11	Kroeker et al. show: A speech recognition system according to claim 2, wherein the coincidence output (output of Fig.8: 228 & Fig.9: 234) includes a sum of products (sum self product and sum cross product) of novelty output points (Fig.8: 228 & Fig.9: 234) of over two sets of novelty output points (e.g., different points of $x_{i,j,n}$ **222**).
Claim(s) 12	Kroeker et al. show: A speech recognition system according to claim 11, wherein the two sets of DFT

	points ($x_{i,j,n}$, $x_{i+\Gamma,j+\Delta,n}$, Fig.9: 234) includes a first set of novelty output points ($x_{i,j,n}$) corresponding to a first instant in time (j) and a second set of novelty output points ($x_{i+\Gamma,j+\Delta,n}$) corresponding to a second time instance (j+Δ). *{The first time instance j is different from the second time instance j+Δ.}*
Claim(s) 13	Kroeker et al. show: A speech recognition system according to claim 11, wherein the two sets of novelty output points ($x_{i,j,n}$, $x_{i+\Gamma,j,n}$, Fig.8: 228) all correspond to a single time instance (j).
Claim(s) 14	Kroeker et al. show: A speech recognition system according to claim 11, wherein the coincidence processor performs the sum of products of novelty output points over two sets of novelty output points (Fig.8: 228 & Fig.9: 234) according to one or more selectably variable coincidence parameters (e.g., delta time, j = 0…5; delta frequency, i = 1…20-Γ, Fig.8: 228) including time duration, frequency extent, base time, base frequency, delta time, delta frequency, and combinations thereof.
Claim(s) 15	Kroeker et al. show: A speech recognition system according to claim 2, wherein each of the time instances further includes an energy value (Fig.3: 112) in addition to the series of DFT points. (col.7, ll.56-60)
Claim(s) 16	Kroeker et al. show: A speech recognition system according to claim 15, wherein the attention processor (see Fig.5) (i) compares the energy value (r_m **132**) to a predetermined threshold value (e.g., value = 21) according to a comparison criterion (Fig.5: 134), so as to produce an energy threshold determination, and (ii) produces the gating signal (s_m = 1) as a predetermined function of the threshold determination (when $r_m \geq 21$).
Claim(s) 17	Kroeker et al. show:

	A speech recognition system according to claim 16, wherein the one or more attention parameters (Fig.3: 134) include the predetermined threshold value (e.g., value = 21), the comparison criterion (Fig.3: 134) and the predetermined function of the threshold determination (e.g., $s_m = 1$ when $r_m \geq 21$).
Claim(s) 19	Kroeker et al. show: A speech recognition system (see Abstract) for transforming a short-time frequency representation (Fig.2: 18; Fig.3) of an acoustic signal (SPEECH, Fig.2) into a stream of coincidence vectors (output of Fig.2: 28; see Fig.8 & 9), comprising: a novelty processor (receptive field processor and adaptive field normalizer, Fig.2: 24, 26; Fig.6, 7) for receiving the short-time frequency representation (e.g., q_m, Fig.6: 130) of the acoustic signal, separating one or more background components (e.g., breathing noises, unvoiced phonemes, col.10, ll.23-27) of the representation from one or more region-of-interest components (e.g., the voice components, col.10, ll.15-16) of the representation, and producing a novelty output (X_n, Fig.7: 222) including the region of interest components (e.g., the voice components) of the representation according to one or more novelty parameters (parameters of the vector w_n, see Fig.7: 214 & col.10, ll.6-7; parameters p, t_n, and voice threshold, see Fig.7: 216 & col.10, ll.37-54); (see col.9, ll.59-68; col.10, ll.1-63) *{1. Fig.6 is a block diagram depicting the receptive field processor of Fig.2: 24; Fig.7 is a block diagram depicting the adaptive normalizer of Fig.2: 26.* *2. The vectors e_m **114**, f_m **116**, q_m **130**, and the matrix V_n **210**, all are short-time frequency representations of the acoustic signal through different stages of processing. For example, q_m, is the output of the inputs e_m and f_m through the intermediate steps of Fig.4 (see col.7, ll.63-68; col.8, ll.1-34). V_n is the result of processing q_m through the steps of Fig.6 and Fig.7: 206 and 208 (see col.9, ll.17-58).* *3. V_n **210** data correspond to a SPEECH signal segment with a significant presence of the voice components (col.9, ll.59-59-62). If the integrated energy, t_n, does not exceed the voice threshold value, e.g., 25, then the adaptive average vector x'_n **218**, which corresponds to the noise components in this case, is subtracted from V_n to produce the matrix X_n, i.e., the voice components (col.10, ll.17-29, 55-63; see Fig.7).}* a coincidence processor (receptive field nonlinear processor, Fig.2: 28; Fig.8, 9) for receiving the novelty output (X_n **222**) and producing a coincidence vector (e.g., output of

	Fig.8: 228 & Fig.9: 234) that includes co-occurrences (e.g., correlation) between samples of the novelty output over time and frequency, (col.12, ll.1-13, 35-46) according to one or more coincidence parameters (e.g., delta time, j = 0…5; delta frequency, i = 1…20-Γ, Fig.8: 228); Kroeker et al. do not show: a coincidence processor for receiving the gating signal. However, Rabiner et al. teach: producing a gating signal (a windowed signal, $\hat{w}_1(m)$ and $\hat{w}_2(m)$, p.147, eq.4.35a, 4.35b) for a coincidence processor (e.g., a processor for calculating the correlation function, eq.4.33, p.146). It would have been obvious to a person of ordinary skill in the art at the time the invention was made to modify the speech recognition system of Kroeker et al. to include the windowing (i.e., gating signal) technique for calculating correlation functions (e.g, the modified short-time correlation function) as taught by Rabiner et al. in order to generate a selectively gated coincidence output. The benefit gain would be that more correlation peaks are displayed at the gated coincidence output (Rabiner et al., p.148, 2nd ¶, ll.6-7) which is useful for determining the periodicity of the speech signal.
Claim(s) 20	Kroeker et al. show: A speech recognition system according to claim 19, further including an attention processor (energy detect processor, Fig.2: 22; Fig.5) for producing a gating signal (s_m, Fig.5: 134) according to one or more attention parameters (time parameter m and s_m values: 0, 1, Fig.5: 134), wherein the coincidence output is produced according to one or more coincidence parameters (e.g., delta time, j = 0…5; delta frequency, i = 1…20-Γ, Fig.8: 228); Kroeker et al. do not show: an attention processor for receiving the novelty output and producing a gating signal as a predetermined function of the novelty output.

	However, Rabiner et al. teach: receiving the novelty output (e.g., different speech samples (x(n+m), x(n+m+k), eq.4.33, p.146) and producing a gating signal (a windowed signal, $\hat{w}_1(m)$ and $\hat{w}_2(m)$, p.147, eq.4.35a, 4.35b) as a predetermined function of the novelty output (e.g., depending on the finite length N of the speech samples) according to one or more attention parameters (e.g., the time parameter m, the window values: 0, 1). (p.146-148) *{1. The gating signal is $\hat{w}_1(m)\hat{w}_2(m+k)$ when applying to the cross-correlation equation of 4.33, which can be written as eq.4.36. The gating signal is a function of time. Eq.4.36 is the cross-correlation function for two different finite length segments of speech (p.148, 2nd ¶).* *2. The finite length N, e.g., N = 6, would correspond to the time unit j = 0…5 of Fig.9: 234.}* It would have been obvious to a person of ordinary skill in the art at the time the invention was made to modify the speech recognition system of Kroeker et al. to include the windowing (i.e., gating signal) technique for calculating correlation functions (e.g, the modified short-time correlation function) as taught by Rabiner et al. in order to generate a selectively gated coincidence output. The benefit gain would be that more correlation peaks are displayed at the gated coincidence output (Rabiner et al., p.148, 2nd ¶, ll.6-7) which is useful for determining the periodicity of the speech signal.
Claim(s) 22	Kroeker et al. show: A method of transforming an acoustic signal into a stream of phonetic estimates, comprising: (see Abstract) receiving the acoustic signal (speech signal s(t), Fig.3: 100) and producing a short-time frequency representation (e_m, Fig.3: 114) of the acoustic signal; (Fig.2: 18; Fig.3; col.7, ll.21-62) *{1. The discrete Fourier transform (DFT) of the finite length vector c_m **108**, i.e., the 128-point DFT vector d_m **110**, is a short-time Fourier transform of the acoustic signal.* *2. The energy vector e_m **114** represents the energy spectrum in the frequency domain, i.e., the short-time frequency representation.}* separating one or more background components (e.g., breathing noises, unvoiced phonemes, col.10, ll.23-27) of the representation from one or more region of interest

components (e.g., the voice components, col.10, ll.15-16) of the representation, and producing a novelty output (X_n, Fig.7: 222) including the region of interest components (e.g., the voice components) of the representation according to one or more novelty parameters (parameters of the vector w_n, see Fig.7: 214 & col.10, ll.6-7; parameters p, t_n, and voice threshold, see Fig.7: 216 & col.10, ll.37-54); (see col.9, ll.59-68; col.10, ll.1-63)

{1. Fig.6 is a block diagram depicting the receptive field processor of Fig.2: 24; Fig.7 is a block diagram depicting the adaptive normalizer of Fig.2: 26.

*2. The vectors e_m **114**, f_m **116**, q_m **130**, and the matrix V_n **210**, all are short-time frequency representations of the acoustic signal through different stages of processing. For example, q_m, is the output of the inputs e_m and f_m through the intermediate steps of Fig.4 (see col.7, ll.63-68; col.8, ll.1-34). V_n is the result of processing q_m through the steps of Fig.6 and Fig.7: 206 and 208 (see col.9, ll.17-58).*

*3. V_n **210** data correspond to a SPEECH signal segment with a significant presence of the voice components (col.9, ll.59-59-62). If the integrated energy, t_n, does not exceed the voice threshold value, e.g., 25, then the adaptive average vector x'_n **218**, which corresponds to the noise components in this case, is subtracted from V_n to produce the matrix X_n, i.e., the voice components (col.10, ll.17-29, 55-63; see Fig.7).}*

producing a coincidence output (e.g., output of Fig.8: 228 & Fig.9: 234) that includes correlations between samples of the novelty output over time and frequency (Fig.8: 228 & Fig.9: 234; col.12, ll.1-13, 35-46), wherein the coincidence output is produced according to one or more coincidence parameters (e.g., delta time, j = 0...5; delta frequency, i = 1...20-Γ, Fig.8: 228);

Kroeker et al. do not show:

producing a gating signal as a predetermined function of the novelty output according to one or more attention parameters;

However, Rabiner et al. teach:

receiving the novelty output (e.g., different speech samples (x(n+m), x(n+m+k), eq.4.33, p.146) and producing a gating signal (a windowed signal, $\hat{w}_1(m)$ and $\hat{w}_2(m)$, p.147, eq.4.35a, 4.35b) as a predetermined function of the novelty output (e.g., depending on the finite length N of the speech samples) according to one or more attention parameters (e.g.,

	the time parameter m, the window values: 0, 1). (p.146-148) *{1. The gating signal is $\hat{w}_1(m)\hat{w}_2(m+k)$ when applying to the cross-correlation equation of 4.33, which can be written as eq.4.36. The gating signal is a function of time. Eq.4.36 is the cross-correlation function for two different finite length segments of speech (p.148, 2nd ¶).* *2. The finite length N, e.g., N = 6, would correspond to the time unit j = 0…5 of Fig.9: 234.}* It would have been obvious to a person of ordinary skill in the art at the time the invention was made to modify the speech recognition system of Kroeker et al. to include the windowing (i.e., gating signal) technique for calculating correlation functions (e.g, the modified short-time correlation function) as taught by Rabiner et al. in order to generate a selectively gated coincidence output and producea phonetic estimate stream representative of acoustic signal as a function of the gated coincidence output. The benefit gain would be that more correlation peaks are displayed at the gated coincidence output (Rabiner et al., p.148, 2nd ¶, ll.6-7) which is useful for determining the periodicity of the speech signal.
Claim(s) 23	Kroeker et al. show: A method according to claim 22, further including (i) calculates a first average value (V_n Fig.7: 210; average by two in time; Fig.6: 204) across a first predetermined frequency range (0…20, Fig.6: 204) and a first predetermined time span (m…m-11, Fig.6: 204), (col.9, ll.17-25) *{The matrix U_n **206** is the result of averaging which becomes V_n **210** after the step of **208**.}* (ii) calculates a second average value (average over time, Fig.7: 212 and accumulative adaptive average, Fig.7: 216) across a second predetermined frequency range (0…20, Fig.7: 214) and a second predetermined time span (0…5, Fig.7: 212), (col.9, ll.59-68; col.10, ll.1-16) and *{The result of second average is the vector $x'_{k,n}$ **218**.}* (iii) subtracts (Fig.7: 220) the second average value ($x'_{k,n}$ **218**) from the first average value (V_n **210**) so as to produce the novelty output point (X_n **222**). (col.10, ll.55-63)
Claim(s) 26	Kroeker et al. show:

	A method according to claim 22, further including comparing the energy value (r_m **132**) to a predetermined threshold value (e.g., value = 21) according to a comparison criterion (Fig.5: 134), so as to produce an energy threshold determination, and (ii) producing the gating signal ($s_m = 1$) as a predetermined function of the threshold determination (when $r_m \geq 21$).

Allowable Subject Matter

7. Claims 4, 5, 10, 18, 21, 24, 25, and 27 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

8. The following is a statement of reasons for the indication of allowable subject matter:

Claim(s) 4	The prior art fails to show: A speech recognition system according to claim 3, wherein the first frequency range, the first time span, the second frequency range and the second time span are each a function of one or more of the novelty parameters.
Claim(s) 5	The prior art fails to show: the first predetermined frequency range is substantially centered about a frequency corresponding to DFT point, and the first predetermined time span is substantially centered about an instant in time corresponding to the DFT point.
Claim(s) 10	The prior art fails to show: A speech recognition system according to claim 3, wherein for each DFT point, the novelty processor further calculates one or more additional novelty outputs, and each additional novelty output is defined by characteristics including a distinct first frequency range, first time span, second frequency range and second time span, each characteristic being a function of one or more of the novelty parameters.
Claim(s) 18	The prior art fails to show: A speech recognition system according to claim 1, wherein the novelty parameters,

	the attention parameters and the coincidence parameters are selected via a genetic algorithm.
Claim(s) 21	The prior art fails to show: A speech recognition system according to claim 19, wherein the novelty parameters and the coincidence parameters are selected via a genetic algorithm.
Claim(s) 24	The prior art fails to show: A method according to claim 22, further including calculating, for each of a plurality of DFT points from the a short-time frequency representation of the acoustic signal, one or more additional novelty outputs, wherein each additional novelty output is defined by characteristics including a distinct first frequency range, first time span, second frequency range and second time span, each characteristic being a function of one or more of the novelty parameters.
Claim(s) 25	The prior art fails to show: A method according to claim 24, further including performing a sum of products of novelty outputs over two sets of novelty outputs according to one or more selectably variable coincidence parameters including time duration, frequency extent, base time, base frequency, delta time, delta frequency, and combinations thereof.
Claim(s) 27	The prior art fails to show: A method according to claim 22, further including selecting the novelty parameters, the attention parameters and the coincidence parameters via a genetic algorithm.

Conclusion

9. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

U.S. Patent Documents:

[1] 2003/0055639 A1 03/2003 Rees

[2] 6,311,153 10/2001 Nakatoh et al.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Tim Lao whose telephone number is 703-305-8955. The examiner can normally be reached on M-F, 8:30am-5pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Doris To can be reached on 703-305-4827. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Tim Lao
Examiner
Art Unit 2655

TL
02/19/04

DORIS H. TO 2/23/04
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2600